УДК 51: 621.72

А. В. Лысенко

ПОСТРОЕНИЕ ЦИФРОВОГО ВОДЯНОГО ЗНАКА, УСТОЙЧИВОГО К ПОТЕРЕ СИНХРОНИЗАЦИИ

1. Введение

1.1. Цифровые водяные знаки и проблема потери синхронизации

Цифровая стеганография – наука о скрытной передаче информации в цифровом виде. Если криптография позволяет защитить содержимое сообщения, то стеганография скрывает само присутствие сообщения: оно незаметно встраивается в сигнал-контейнер, который в дальнейшем передается по открытому каналу. В качестве контейнера может использоваться цифровое изображение, оцифрованный звуковой или видеосигнал.

В классической задаче стеганографии контейнер вторичен по отношению к информации, содержащейся в нем, и служит лишь прикрытием. Однако существуют разделы стеганографии, где приоритеты расставлены иначе. Один из таких разделов – внедрение цифровых водяных знаков.

Понятие цифровых водяных знаков (ЦВЗ) возникло в период повышенного интереса к проблеме защиты авторских прав на информацию, распространяемую в цифровом виде. Объектом нарушения авторских прав может стать любое изображение, аудио- или видеофайл, выложенный автором в сети Интернет или распространяемый по иным каналам. Чтобы облегчить обнаружение факта несанкционированного использования авторских материалов, а также обеспечить возможность доказательства авторских прав, перед опубликованием документа в доступном широкой публике виде в него с помощью стеганографической техники внедряется дополнительная информация – авторская метка, или цифровой водяной знак.

В отличие от классической постановки задачи стеганографии цифровой водяной знак внедряется для того, чтобы защитить контейнер. Изменения, возникающие в контейнере при внедрении водяного знака, должны быть минимальны.

Важным условием успешного использования цифрового водяного знака является его устойчивость по отношению к атакам – попыткам злоумышленника разрушить водяной знак или сделать его нечитаемым. Типичными атаками в случае цифрового изображения являются изменение формата,

сжатие с потерей качества, обрезание краев или вырезание фрагмента, масштабирование, поворот, фильтрация и добавление шума. Если после одной или нескольких атак контейнер для восприятия изменился несущественно, а водяной знак не может быть найден алгоритмом обнаружения, то цель злоумышленника достигнута.

Часто цифровой водяной знак не может быть обнаружен после атаки вследствие *проблемы потери синхронизации*. Опишем эту проблему.

Рассмотрим в качестве примера простейший способ внедрения информации в младший бит яркости изображения. Изображение в градациях серого, яркость каждого пикселя которого записана числом от 0 до 255, представляется построчно в виде одномерного массива байт, после чего младший бит каждого байта заменяется на очередной бит внедряемой информации.

При отсутствии атак, представив изображение в виде одномерного массива, можно считать внедренную информацию из младших бит. Используя большее число бит и коды с исправлением ошибок, можно повышать устойчивость этой схемы к сжатию с потерей качества, фильтрации и зашумлению. Однако если злоумышленник вырежет фрагмент изображения, повернет его на небольшой угол или изменит масштаб, обнаружение водяного знака станет невозможным, так как мы не сможем восстановить исходный одномерный массив в атакованном изображении, не зная, каким трансформациям оно подвергалось. Таким образом, знак не разрушен и по-прежнему присутствует в изображении, но не может быть считан из-за *потери синхронизации*.

Существует несколько подходов к решению проблемы потери синхронизации.

Во-первых, можно внедрять водяной знак в область, инвариантную к заданным преобразованиям. В случае аффинных преобразований и вырезания фрагмента эффективно внедрение знака в область двумерного преобразования Фурье [1, 2].

Во-вторых, при внедрении можно добавить сигнал-пилот в частотную область изображения, чтобы обратить аффинные преобразования изображения [3].

Но наиболее перспективными в отношении устойчивости к атакам, вызывающим потерю синхронизации, представляются методы внедрения водяных знаков, использующие для борьбы с рассинхронизацией инвариантные по отношению к атакам характеристики изображения. К таким характеристикам можно отнести информацию о количестве и расположении заметных точек изображения (например, угловых [4]) или специальным образом выбранные характеристики окрестностей точек [5].

Важное преимущество использования инвариантных характеристик изображения состоит в том, что их изменение для введения в заблуждение ал-

горитма обнаружения, как правило, трудоемко и с большой вероятностью повлечет за собой существенные для восприятия изменения в изображении.

1.2. Псевдослучайное бинарное разбиение изображения

При описании алгоритмов внедрения цифровых водяных знаков часто используется понятие *псевдослучайного бинарного разбиения изображения*. Это понятие понадобится нам в дальнейшем.

Псевдослучайным бинарным разбиением изображения называется отображение $h(x, y)$ пикселей изображения в двухэлементное множество $\{0, 1\}$, которое может быть построено с помощью ключевой информации K. При этом поведение функции отображения для стороннего наблюдателя, не обладающего ключевой информацией, неотличимо от поведения случайной функции, равновероятно принимающей значения 0 и 1 на пикселях изображения.

Псевдослучайное бинарное разбиение изображения можно задать, в частности, с помощью псевдослучайной бинарной функции $f(i) : \mathbb{N} \to \{0, 1\}$, построенной с помощью ключа K. Зададим разбиение на изображении следующим образом: сопоставим каждому пикселю (x, y) изображения его порядковый номер n в массиве построчного представления изображения и определим функцию разбиения на нем: $h(x, y) = f(n)$.

Определив функцию псевдослучайного разбиения изображения, можно говорить о том, что изображение разбито на две части: те пиксели, которые переводятся этой функцией в 0, и те, которые переводятся в 1. Псевдослучайное разбиение изображения можно использовать, например, чтобы внедрить один бит информации в изображение.

Произведем следующую операцию: добавим небольшое число d к яркости пикселей одной части изображения (пикселей, которые отображаются в 1 псевдослучайным разбиением) и вычтем то же число d из яркости пикселей второй части изображения (пикселей, которые отображаются в 0). Если в исходном изображении средние яркости первой и второй части изображения примерно совпадают, то после описанной операции средние яркости первой и второй частей изображения будут различаться на $2d$ – величину, существенно отличную от нуля. Будем считать, что внедрен бит $<1>$, если $d > 0$, и $<0>$, если $d < 0$.

Информационный бит, внедренный таким образом, устойчив к сжатию с потерей качества, добавлению случайного шума и фильтрации изображения. Устойчивость же его по отношению к описанным выше атакам, которые направлены на потерю синхронизации, определяется устойчивостью используемого псевдослучайного разбиения по отношению к этим атакам. Легко видеть, что разбиение в приведенном примере такой устойчивостью обладать не будет, так как принадлежность пикселя к соответствующей части определяет-

ся на основе координаты пикселя в построчном представлении изображения, которое может сильно меняться после атаки.

Таким образом, очевидна важность методов построения устойчивого псевдослучайного бинарного разбиения изображения. Один из новых результатов в этой области представлен Дамьеном Деланнэ в [5]. В следующих разделах мы рассмотрим этот метод и предложим некоторые усовершенствования процесса построения разбиения.

2. Алгоритм Деланнэ псевдослучайного разбиения изображения

Отметим следующий факт: для построения устойчивого псевдослучайного бинарного разбиения принадлежность пикселя к той или иной части разбиения должна определяться из характера изображения в некоторой окрестности пикселя изображения. Эта окрестность должна быть достаточно маленькой, чтобы получить устойчивость к вырезанию части изображения и другим атакам, но и достаточно большой, чтобы можно было построить устойчивую нетривиальную псевдослучайную функцию, основываясь на ней. Первый шаг алгоритма Деланнэ состоит в нахождении характеристического радиуса окрестности для каждой точки изображения. Данный характеристический радиус должен быть устойчив к геометрическим трансформациям изображения. Это значит, что для фиксированной точки изображения после поворота изображения радиус не должен измениться, а при изменении масштаба изображения он должен измениться пропорционально изменению масштаба изображения.

Деланнэ предложил искать характеристический радиус, расширяя круговую окрестность точки до выполнения определенного условия. Обозначим среднюю яркость диска с центром в точке (x, y) радиуса r как $M_{x,y,disk}(r)$:

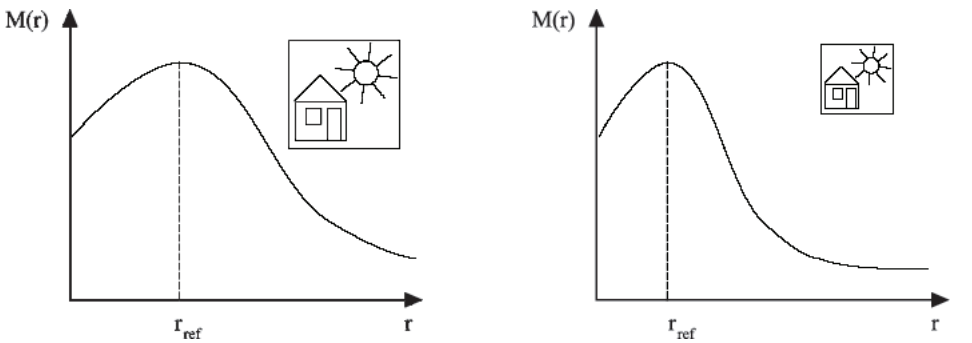
$$M_{x,y,disk}(r) = \frac{1}{\pi r^2} \int_0^r \int_0^{2\pi} I(x + \rho \cdot \cos\theta, y + \rho \cdot \sin\theta)\, d\theta d\rho. \tag{1}$$

Деланнэ предлагает найти такой радиус r_{ref}, что производная средней яркости диска по r будет равна нулю в точке r_{ref}.

$$r_{ref}(x, y) = \min \left\{ r \in \mathbb{R} \;\middle|\; \frac{dM_{x,y,disk}}{dr}(r) = 0 \right\}. \tag{2}$$

Легко видеть, что характеристический радиус, определенный таким образом, будет удовлетворять требованиям устойчивости к повороту и изменению масштаба изображения, изложенным выше. Устойчивость к повороту обусловлена использованием круговой окрестности, а устойчивость к масштабированию вытекает из того факта, что функция $M_{x,y,disk}(r)$ для фиксирован-

ной точки при растяжении/сжатии изображения растягивается/сжимается соответственно (рис. 1). Использование низкочастотных характеристик изображения (средняя яркость диска) позволяет также ожидать хорошей стойкости полученного характеристического радиуса к высокочастотным изменениям изображения, таким как добавление шума, фильтрация или сжатие с потерей качества.



а *б*

Рис. 1. Средняя яркость диска радиуса r: *а* – для выбранного пикселя исходного изображения; *б* – для соответствующего пикселя отмасштабированного изображения

Более подробно процедура построения характеристического радиуса будет обсуждаться в следующей части.

С помощью построенного характеристического радиуса алгоритм определяет принадлежность точки к той или иной части разбиения. Деланнэ предлагает выбрать два ключа $k1$ и $k2$ и рассмотреть две окружности радиусов

$$r_1 = k1 \cdot r_{ref}, \qquad r_2 = k2 \cdot r_{ref}.$$

Далее вычисляются две кривые изменения яркости C_1 и C_2 вдоль кругов радиусов r_1 и r_2 соответственно:

$$C_1(\theta) = i(r_1, \theta), \qquad C_2(\theta) = i(r_2, \theta).$$

Кривая C_t определяется как сумма C_1 и C_2:

$$C_t(\theta) = C_1(\theta) + C_2(\theta).$$

Принадлежность пикселя к части бинарного разбиения определяется с помощью углового расстояния α между максимальным и минимальным значением $C_t(\theta)$: $h(x, y) = 1$, если $\alpha \geq \pi$, и $h(x, y) = 0$, если $(\alpha < \pi)$.

Деланнэ также обсуждает альтернативные функции принадлежности и вопросы секретности полученного разбиения. Желающие могут ознакомиться с ними в [5]. Мы же вернемся к базовой процедуре алгоритма – нахождению характеристического радиуса и предложим ряд усовершенствований в ней.

3. Вариационный характеристический радиус

Первая проблема в алгоритме нахождения характеристического радиуса, предложенном Деланнэ, возникает при вычислении средней яркости для дисков малого радиуса. Такие вычисления используют значения яркости небольшого числа соседних пикселей изображения и поэтому весьма неточны и чувствительны к высокочастотным изменениям изображения, таким, например, как добавление шума.

Чтобы преодолеть эту проблему, Деланнэ предлагает использовать порог r_{min} величиной порядка 5 пикселей.

Характеристический радиус с учетом порога r_{min} определяется следующим образом:

$$r_{ref}(x,y) = \arg \min_{r \in [r_{min}, r_{max}]} \frac{dM_{x,y,disk}}{dr}(r). \tag{3}$$

Использование порога r_{min} позволяет получить большую устойчивость алгоритма к высокочастотным изменениям изображения, но также привносит в алгоритм неустойчивый к масштабированию параметр r_{min}.

Вторая проблема тесно связана с первой. В однородных (или почти однородных) областях изображения минимальное значение производной часто может достигаться на минимальном допустимом значении радиуса $r_{ref} = r_{min}$. Так как r_{min} – постоянная алгоритма, то после масштабирования изображения с фактором Sc и вычисления характеристического радиуса мы получим $r_{ref} = r_{min}$ вместо $r_{ref} = Sc \cdot r_{min}$. Это может стать серьезной проблемой, так как однородные области могут составлять значительную часть изображения.

Деланнэ также рассматривает характеристический радиус *пика B*, который находится как решение (2), ближайшее к заданному значению r_{tg}. Как бы то ни было, это значение также неустойчиво к масштабированию, что приводит к сходным проблемам при вычислении характеристического радиуса.

Следствием описанных проблем является ограниченная устойчивость характеристического радиуса по отношению к масштабированию изображения.

Чтобы преодолеть описанные сложности, мы предлагаем альтернативную процедуру нахождения характеристического радиуса.

Так же как и в исходном алгоритме, мы будем искать характеристический радиус, расширяя окрестность вокруг выбранной точки до тех пор, пока не будет выполнено определенное условие. По возможности будем искать такое

условие, которое не привлекало бы искусственных порогов, таких как r_{min} или r_{tg}.

Основная идея заключается в том, чтобы построить характеристику диска изображения, которая 1) монотонно возрастает с ростом размера диска и 2) независима от масштаба, т. е. данная характеристика должна оставаться почти неизменной для выбранной области изображения и соответствующей области отмасштабированного изображения.

В качестве такой характеристики можно рассматривать количество визуально значимой информации, заключенной в области изображения, или *меру сложности* этой области. Алгоритм нахождения характеристического радиуса изображения тогда будет сводиться к расширению диска до тех пор, пока характеристика сложности области изображения, заключенной в этом диске не достигнет заданного порога. Преимущество такого подхода заключается в том, что радиус, удовлетворяющий такому условию, всегда единственный (следствие монотонности функции сложности), а значит, решение будет более устойчивым. Второе преимущество заключается в том, что мы можем варьировать порог необходимой минимальной сложности области, таким образом контролируя размер и устойчивость характеристического радиуса. Проблема, связанная с нахождением характеристического радиуса для однородных областей изображения, также решается – мера сложности изображения в таких областях будет мала, и радиус диска будет расширяться, пока в него не попадет достаточное количество визуально значимых деталей изображения.

Мы рассмотрели несколько определений меры сложности области изображения. В одномерном случае в качестве меры сложности функции можно использовать ее вариацию

$$v(f,a,b)=\sum_{x=a}^{b-1}|f(x+1)-f(x)|. \tag{4}$$

По аналогии мы определили двумерную вариацию диска изображения как

$$V(I,x_0,y_0,r)=\sum_{(x,y)\in disk(x_0,y_0,r)}(|I(x+1,y)-I(x,y)|+|I(x,y+1)-I(x,y)|), \tag{5}$$

где $I(x,y)$ — значение яркости пикселя с координатами (x,y).

Эксперименты показали, что аналог одномерной вариации $V(I,x_0,y_0,r)$ не обладает достаточной устойчивостью к масштабированию изображения, но хорошая устойчивость к масштабированию была достигнута при использовании характеристики

$$W(I,x_0,y_0,r)=\frac{V^2(I,x_0,y_0,r)}{r^2}. \tag{6}$$

Она незначительно возрастает, если изображение растягивается (действительно, в этом случае мы имеем почти то же количество информации в выбранной области изображения), и становится немного меньше, если изображение сжимается (в этом случае мы теряем некоторую часть информации, содержащуюся в тонких деталях изображения). Таким образом, характеристика $W(I, x_0, y_0, r)$ может использоваться в качестве меры сложности изображения.

Однако, несмотря на хорошую устойчивость характеристики W к геометрическим трансформациям, она не подходит для внедрения устойчивого водяного знака из-за своей хрупкости по отношению к добавлению шума к изображению. Действительно, W получается суммированием разностей значений яркости соседних пикселей, и добавление шума даже малой амплитуды к изображению I значительно увеличивает значение $W(I, x, y, r)$ для выбранной области.

Для преодоления этой проблемы мы предложили использовать метод вычисления вариации совместно с подходом Деланнэ. Вместо анализа сложности двумерного диска изображения будем анализировать сложность одномерной функции средней яркости диска изображения в зависимости от радиуса $M_{x,y,disk}(r)$ (1). В качестве меры сложности одномерной функции будем использовать стандартную одномерную вариацию v (4). Таким образом, для нахождения характеристического радиуса изображения мы будем расширять область вокруг пикселя, пока вариация функции $M_{x,y,disk}(r)$ на отрезке $[0, r]$ не достигнет заданного значения C:

$$r_{ref}(x, y) = \min \left\{ r \in \mathbb{R} \,\middle|\, v(M_{x,y,disk}, 0, r) = C \right\}. \tag{7}$$

Договоримся в дальнейшем называть характеристический радиус, определенный таким образом, *вариационным*, а характеристический радиус, найденный по алгоритму Деланнэ, *экстремальным* или *характеристическим радиусом Деланнэ*.

Эксперименты показали хорошую устойчивость вариационного характеристического радиуса к повороту и масштабированию с фактором, большим единицы, и удовлетворительную устойчивость к масштабированию с фактором, меньшим единицы. К сожалению, устойчивость характеристического радиуса по отношению к добавлению шума ограничена. Тем не менее влияние высокочастотных изменений изображения может быть скомпенсировано увеличением шага радиуса до нескольких пикселей при вычислении вариации функции $M_{x,y,disk}(r)$.

Процедура вычисления вариационного радиуса имеет ту же вычислительную сложность, что и процедура вычисления экстремального характеристи-

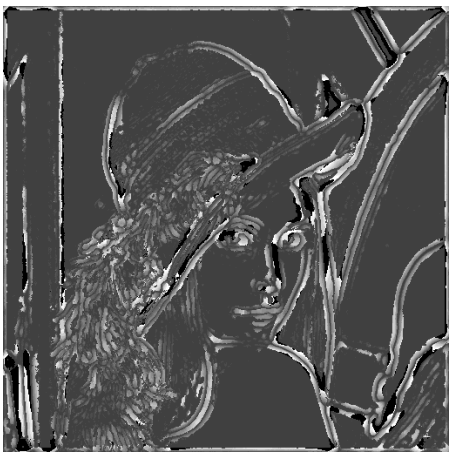
ческого радиуса Деланнэ, и может быть эффективно реализована с помощью последовательности сверток изображения с дисковыми фильтрами. Более того, оба радиуса могут быть найдены одновременно за один проход алгоритма, что позволяет дать, например, следующие гибридные определения характеристического радиуса.

Характеристический радиус r_{ref} – это решение задачи (2) с вариацией $v(M_{x,y,disk}, 0, r_{ref})$, превосходящей заданный порог C (или ближайшей к нему). Иначе говоря, мы можем заменить константы r_{min} или r_{tg} в алгоритме Деланнэ на устойчивый к геометрическим преобразованиям вариационный характеристический радиус (7).

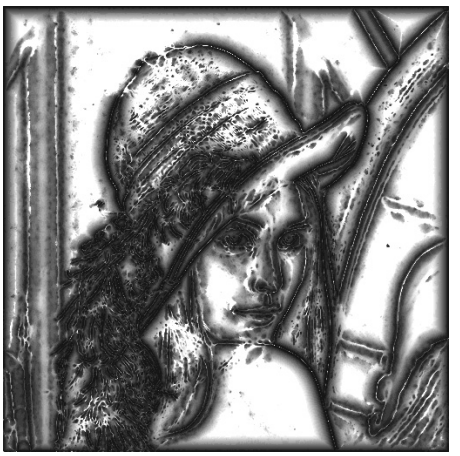
Приведем пример вычисления характеристического радиуса Деланнэ и вариационного характеристического радиуса для изображения Лены (рис. 2).

Рис. 2. Изображение Лены

На рис. 3 показаны значения экстремального и вариационного характеристических радиусов изображения. Более яркие пиксели соответствуют большим значениям характеристического радиуса. Видно, что однородные области изображения имеют малый экстремальный характеристический радиус и большой вариационный, что хорошо согласуется с выкладками, сделанными выше. Для областей изображения, соответствующих тонким деталям, характерно обратное положение вещей: экстремальный характеристический радиус, как правило, больше вариационного. Такое поведение характеристических радиусов наводит на мысль об эффективном взаимном дополнении определений и целесообразности использования гибридных определений характеристического радиуса.



а



б

Рис. 3. Экстремальный (*а*) и вариационный (*б*) характеристические радиусы изображения

3.1. Фильтр адаптивного размытия изображения

Интересным побочным продуктом вычисления вариационного характеристического радиуса является фильтр адаптивного размытия изображения, действующий следующим образом: значение яркости каждого пикселя заменяется на среднее значение яркости диска с центром в этом пикселе и радиусом, равным характеристическому радиусу изображения в этой точке. Поскольку характеристический радиус для областей, содержащих тонкие детали, мал и, напротив, велик для однородных областей изображения, адаптивное размытие позволяет размыть изображение, сохранив при этом границы и тонкие детали (рис. 4). Напомним, что обычно фильтры размытия реализуются с помощью свертки изображения с изображением–фильтром фиксированного радиуса. Желающие могут получить необходимую информацию по фильтрации и обработке изображений в [6].

Фильтр адаптивного размытия (как и алгоритм нахождения вариационного радиуса) требует последовательности сверток изображения с фильтрами увеличивающегося радиуса. При этом критерий прекращения размытия определяется отдельно для каждой точки изображения.

Степень размытия может регулироваться посредством изменения необходимой минимальной вариации в определении вариационного характеристического радиуса.

Фильтр адаптивного размытия близок к фильтру Smart Blur пакета Adobe Photoshop, но принципиально отличается по принципу действия. Алгоритм

а *б*

Рис. 4. Исходное изображение Лены (*а*) и результат его адаптивного размытия (*б*)

размытия Smart Blur сначала выделяет границы изображения и разбивает изображение на области-кластеры, внутри которых происходит размытие с фиксированным радиусом, сохраняющее четкие границы между кластерами. Таким образом, для работы алгоритма Smart Blur требуется две константы – порог обнаружения границ и величина размытия. Основным недостатком такого подхода является то, что если деталь изображения недостаточно выражена, чтобы быть квалифицированной как «граница», она может быть размыта Smart Blur до полной неразличимости. Понижение же порога обнаружения границ значительно увеличивает сложность алгоритма и приводит к возникновению множества мелких кластеров, что понижает эффективность размытия.

Фильтр адаптивного размытия лишен описанного недостатка, поскольку не разбивает точки на «граничные» и «гладкие», а вычисляет степень сложности изображения каждой точки, в соответствии с которым определяется степень размытия.

4. Заключение

Мы описали проблему потери синхронизации, актуальную при внедрении устойчивых цифровых водяных знаков в изображение, и рассмотрели один из новых результатов в этой области, представленный в [5].

Нами предложен альтернативный вариационный способ определения характеристического радиуса (7), который может быть использован вместо или

совместно с определением Деланнэ (3) для построения устойчивого псевдослучайного разбиения изображения. Использование вариационного характеристического радиуса позволяет решить проблему малых радиусов в алгоритме Деланнэ. Вариационный радиус имеет ряд преимуществ по отношению к атаке масштабирования, но, с другой стороны, он может быть менее устойчив к выскочастотным изменениям изображения.

Разное поведение вариационного и экстремального характеристических радиусов на однородных областях изображения и областях, содержащих тонкие детали, позволяет предложить эффективность гибридных определений характеристического радиуса.

Попутно был получен также способ адаптивного размытия изображения, позволяющий сохранять границы и тонкие детали.

Литература

1. Ruanaidh J. J. K. O., Pun T. Rotation, scale and translation invariant spread spectrum digital image watermarking // Signal Processing. 1998. Vol. 66, № 3. P. 303–318.

2. Lin C., Wu M., Bloom J. A. et al. Rotation, scale, and translation resilient public watermarking for images // Proceedings of the SPIE. 2000. Vol. 3971. P. 90–98.

3. Alvarez-Rodriguez M., Perez-Gonzalez F. Analysis of pilotbased synchronization algorithms for watermarking of still images // Signal Processing: Image Communication. 2002. Vol. 17, № 8. P. 611–633.

4. Bas P., Chassery J.-M., Macq B. Geometrically invariant watermarking using feature points // IEEE Transactions on Image Processing. 2002. Vol. 11, № 9. P. 1014–1028.

5. Delannay D. Digital Watermarking Algorithms Robust Against Loss of Synchronization: Dissertation / Universite catholique de Louvain. Louvain-la-Neuve, 2004.

6. Russ J. C. The Image Processing Handbook (3rd ed.). Boca Raton: CRC Press, 1998.